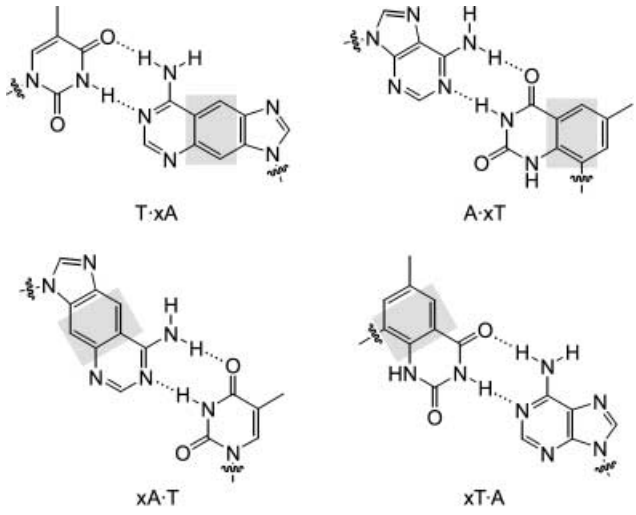
# A New DNA Analogue with Expanded Size and Scope

Alexander Heckel*[a]

In the 50 years since the discovery of the double-helical structure of DNA,[1,2] we have seen a cornucopia of DNA analogues being synthesized and studied. Compared to regular DNA, the modified versions have novel and improved properties, can be used as new diagnostic or architectural tools, and have given us an insight into why and how DNA can act as a storage medium for genetic information and as much more. Virtually all parts of natural nucleic acids—for example the nucleobases—have already been replaced by non-natural modifications.[3] Systematic variation of the hydrogen-bond donor and acceptor positions on the "Watson–Crick face" of the nucleobases afforded an expanded genetic alphabet with twelve replicatable letters[4] that can, in principle, encode for a lot more than the normal twenty amino acids. One of these non-natural nucleobases has already been used in an in vitro translation experiment in which iodotyrosine was built into a polypeptide.[5] Attachment of functional groups to the nucleobases afforded "functionalized DNA" (fDNA);[6] this might lead to nucleic acids with improved catalytic activity. In a different series of studies, the ribose unit was replaced by a six-membered ring (homo-DNA) in order to try to answer the interesting question of why nature chose ribofuranose for the attachment of the nucleobases in DNA.[7] Other projects have been aimed at modifying or completely replacing the sugar phosphate backbone of DNA. These modifications can be as subtle as the replacement of one of the oxygen atoms in the phosphates with sulfur or a methyl group, to yield nuclease-resistant phosphorothioate or methylphosphonate DNA,[8] or as dramatic as the replacement of the entire natural backbone by a peptidic structure to yield "peptide nucleic acids" (PNA).[9,10] Other backbone modifications include morpholino oligonucleotides,[11] locked nucleic acids[12] or bicyclo-DNA;[13] some of these have been successfully used in the antisense approach to gene regulation.[14]

In a recent publication, Kool et al. introduced a new modification in which the nucleobases are size-expanded (Scheme 1).[15] So far only adenine and thymine have been modified: by the introduction of an additional benzene ring, each base has been lengthened by 2.4 Å. These modified nucleobases should still allow for selective recognition of their natural counterparts as shown in Scheme 1. To be able to use the same notations as for unmodified DNA the new size-expanded nucleosides were abbreviated to xA and xT. The phosphoramidates required for automated solid-phase synthesis were accessible in eleven and ten steps, respectively.

To test the compatibility of the new nucleobases with the natural ones, they were embedded in oligonucleotide dodecamers (that otherwise consisted of natural nucleotides) as shown in Figure 1, and the pairing properties have been studied by thermal-denaturation experiments. It turned out that, after exchanging A with xA, a duplex was still formed, but the melting temperature was decreased by about 5 °C (top two examples in Figure 1 with N = T), Kool et al. attributed this to the distortion in the backbone of the duplex that is needed to accommodate the new size-expanded base. Interestingly, xA still showed a very clear pairing preference for T, and pairing of xA with adenine, guanine and cytosine afforded duplexes with significantly lower melting temperatures. Almost the same was true when T was replaced by xT, only in this case it turned out that xT paired equally well with adenine and thymine.




***Scheme 1.*** *Design and putative hydrogen-bonding pattern in the four possible base-pairs between the size-expanded nucleobases xA and xT and their natural counterparts T and A. The boxes highlight the parts where the nucleobases have formally been size-expanded.*

[a] *Dr. A. Heckel*
*Universität Bonn*
*Kekulé-Institut für Organische Chemie*
*und Biochemie*
*Gerhard-Domagk-Straße 1,*
*53121 Bonn (Germany)*
*Fax: (+49) 228-73-5388*
*E-mail: heckel@uni-bonn.de*

 **DOI: 10.1002/cbic.200400001**  

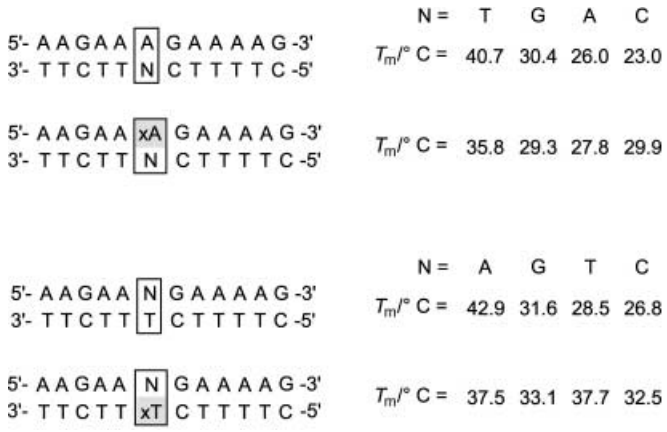



***Figure 1.*** *Results of thermal denaturation experiments with the new size-expanded nucleobases in the context of natural DNA. For experimental details see the supplementary material of the publication by Kool et al.*[15]

If only size-expanded base-pairs are used in the design of a duplex, the concomitant energetic penalties of a mixed-width double helix might be avoided. For this type of size-expanded DNA, the term "xDNA" has been coined by Kool et al. Figure 2 shows a selection of the oligonucleotides synthesized for that investigation. The fact that the modified oligonucleotide strands indeed form a stable structure has been established in different ways. The thermal denaturation curves show an apparent two-state melting behaviour, and the xDNA duplexes turned out to be thermally much more stable (Figure 2); Kool et al. attributed this to enhanced base stacking with the larger π-electron systems of the size-expanded bases. These differences in melting temperature can be as big as 34.3 °C in only a decamer! If all of the size-expanded bases are on one strand, this can serve as a strong binding template for hybridization assays or in antisense applications. Careful examination of the data obtained in thermal denaturation experiments showed evidence for a 1:1 as well as for a 2:1 stoichiometry in certain cases. This behaviour of xDNA—which is also found with natural DNA in certain cases[16]—will have to be examined further, and rules as to when and how this occurs will have to be established. NMR spectra of xDNA showed signals for the hydrogen-bonded imino protons, which would be expected from an antiparallel duplex, and CD spectra of xDNA suggested that the duplex formed has indeed a right-handed helical structure. Figure 3 shows a plausible structure of xDNA as modelled by Kool et al. This structure is in accordance with the data obtained in the NMR experiments. It still has to be interpreted with due caution until more or more precise structural data are available, but several key features can plausibly be expected. With an increase of the diameter of the helix on going from DNA to xDNA, the widths of the major and minor grooves are increased as well. There are both natural and non-natural (macro)molecules that can bind to DNA sequence-specifically, such as transcription factors or small molecules of the pyrrole-imidazole polyamide type.[18] The latter bind to DNA in the minor groove and "read" the hydrogen-bond donor and acceptor patterns. In xDNA, if all the size-expanded bases are on one strand, the hydrogen-bond donor and acceptor patterns could be sufficiently similar to those of unmodified DNA to still allow recognition with the same molecules used for unmodified DNA. However, due to the different point of attachment of the ribose unit in

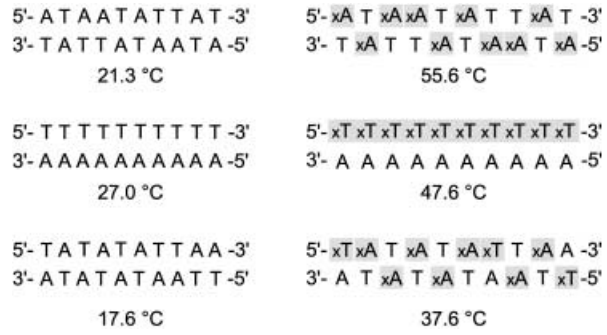



***Figure 2.*** *Melting temperatures of fully size-expanded DNA (xDNA). For experimental details see the supplementary material of the publication by Kool et al.*[15]

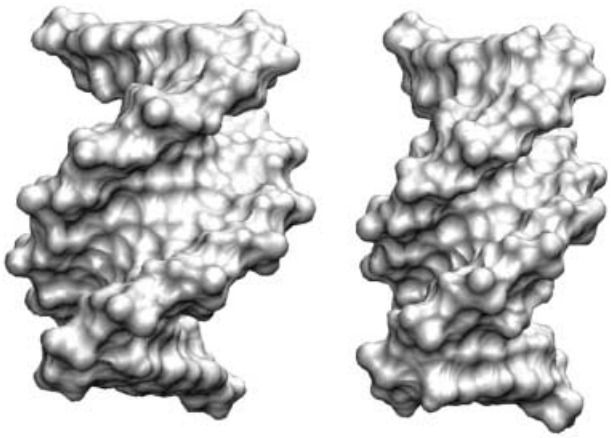


***Figure 3.*** *A plausible modelled structure*[17] *of an xDNA decamer (left) in comparison with a standard B-DNA helix (right). The data files were kindly provided by Kool et al.*

the xT nucleoside, there is an additional hydrogen-bond donor available in the minor groove. If size-expanded bases are present on both strands of a duplex the recognition patterns will probably differ substantially from the ones of unmodified DNA. On the other hand this could make it possible to find DNA-binding agents that can differentiate between all four extended base pairs shown in Scheme 1.

xDNA has at least one more notable feature. The size-expanded bases xA and xT are fluorescent—both as monomers and in xDNA. This might lead to novel applications in detection and analysis of DNA and RNA. It will be interesting to study the fluorescence behaviour under different conditions and the possibility of quenching this fluorescence.

In summary, Kool et al. have shown that the sizes of DNA double helices are not limited to the ones found in nature, and that size-expanded DNA (xDNA) can be more stable than the natural analogue. Having done that, they have opened the door to a new, lateral dimension, with a wealth of opportunities to introduce substitutions and modifications that can lead to new diagnostic, therapeutic or architectural tools. With the size-expanded bases xA and xT it is already possible to form four stable base pairs with their natural counterparts T and A, even though the mismatch (in)-tolerance in the context of xDNA has not yet been investigated. It will be intriguing to follow the completion of the "size-expanded alphabet" with the synthesis of xG and xC; this would provide eight base pairs in conjunction with the natural counterparts. Maybe even a natural or non-natural polymerase can be found that incorporates the size-expanded base pairs and—as mentioned before—it could be possible to find (or evolutionally develop) DNA-binding agents that can bind to xDNA sequence-specifically and differentiate between all possible base-pairs. It might also be interesting to study how far one can go in terms of size-expansion (for example it has not yet been shown if xA and xT alone can form a duplex) before the hydrophobic stacking interactions take over and the recognition properties are lost.

## Acknowledgements

*A.H. gratefully acknowledges financial support from the German Fonds der Chemischen Industrie (Liebig-Stipendium).*

**Keywords:** DNA analogues · nucleic acids · nucleobases · nucleotides

[1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737 – 738.
[2] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 964 – 967.
[3] E. T. Kool, *Chem. Rev.* **1997**, *97*, 1473 – 1487.
[4] S. A. Benner, T. R. Battersby, B. Eschgfäller, D. Hutter, J. T. Kodra, S. Lutz, T. Arslan, D. K. Bäschlin, M. Blättler, M. Egli, C. Hammer, H. A. Held, J. Horlacher, Z. Huang, B. Hydrup, T. F. Jenny, S. C. Jurczyk, M. König, U. v. Krosigk, M. J. Lutz, L. J. MacPherson, S. E. Moroney, E. Müller, K. P. Nambiar, J. A. Piccirilli, C. Y. Switzer, J. J. Vögel, C. Richert, A. L. Roughton, J. Schmidt, K. C. Schneider, J. Stackhouse, *Pure Appl. Chem.* **1998**, *70*, 263 – 266.
[5] J. D. Bain, C. Switzer, A. R. Chamberlin, S. A. Benner, *Nature* **1992**, *356*, 537 – 539.
[6] O. Thum, S. Jäger, M. Famulok, *Angew. Chem.* **2001**, *113*, 4112 – 4115; *Angew. Chem. Int. Ed.* **2001**, *40*, 3990 – 3993.
[7] K. Groebke, J. Hunziker, W. Fraser, L. Peng, U. Diederichsen, K. Zimmermann, A. Holzner, C. Leumann, A. Eschenmoser, *Helv. Chim. Acta* **1998**, *81*, 375 – 474.
[8] S. Verma, F. Eckstein, *Annu. Rev. Biochem.* **1998**, *67*, 99 – 134.
[9] P. E. Nielsen, *Acc. Chem. Res.* **1999**, *32*, 624 – 630.
[10] E. Uhlmann, A. Peyman, G. Breipohl, D. W. Will, *Angew. Chem.* **1998**, *110*, 2954 – 2983; *Angew. Chem. Int. Ed.* **1998**, *37*, 2796 – 2823.
[11] J. Heasman, *Dev. Biol. (San Diego CA U.S.)* **2002**, *243*, 209 – 214.
[12] A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen, J. Wengel, *Tetrahedron* **1998**, *54*, 3607 – 3630.
[13] D.-R. Ahn, A. Egger, C. Lehmann, S. Pitsch, C. J. Leumann, *Chem. Eur. J.* **2002**, *8*, 5312 – 5322.
[14] E. Uhlmann, A. Peyman, *Chem. Rev.* **1990**, *90*, 544 – 584.
[15] H. Liu, J. Gao, S. R. Lynch, Y. D. Saito, L. Maynard, E. T. Kool, *Science* **2003**, *302*, 868 – 871. For a more detailed description of the synthesis and properties of the single nucleobases xA and xT see also: H. Liu, J. Gao, L. Maynard, Y. D. Saito, E. T. Kool, *J. Am. Chem. Soc.* **2004**, *126*, 1102 – 1109.
[16] G. Prakash, E. T. Kool, *J. Am. Chem. Soc.* **1992**, *114*, 3523 – 3527.
[17] Both representations of the duplexes show a Conolly surface with a probe radius of 1.4 Å. The images were generated with VMD 1.8.2 (http://www.ks.uiuc.edu/Research/vmd).
[18] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215 – 2235.

Received: January 5, 2004